`, `Iraq ReCollection: A Proposal for Preserving Iraq’s Cultural Heritage`

`Iraq ReCollection: A Proposal for Recovering Iraq’s Cultural Heritage`

**Statement of significance of the project (Abstract)**

Iraq has a rich tradition of journalistic publishing; its academic journals are invaluable for humanistic and area studies, particularly in history, literature, and politics. Though these journals have been part of a long established creative, scholarly tradition, library holdings in Iraq and globally are limited; electronic catalogs of these holdings are only now beginning to exist and none of the titles has been preserved digitally or is available for widespread viewing. OACIS¹, an evolving electronic union catalog of serials from or about the Middle East developed and maintained by Yale University Library, now provides systematic information for locating Middle Eastern serials in 16 libraries in the United States, Europe, and the Middle East. OACIS identifies approximately 11,000 unique titles representing 88 countries in over 50 languages. Of these titles, close to 600 are published in Iraq, 350 being unique. Approximately half pertain to the Humanities. Yale seeks NEH funding to conduct a digitization project – *Iraq ReCollection* – with the following goals:

- **To digitize** a select group of the most important scholarly humanistic Iraqi journals held by Yale and the University of Pennsylvania, an OACIS partner. This group includes 9 titles. These are published in Arabic; two have articles in both Arabic and Western languages.
- **To create** an electronic archive of these digitized files that permits 1) retrieval and display via the Internet, and 2) integration into other existing electronic systems, such as the search engine of OACIS, so that scholars in Iraq and around the world can easily gain access to this important segment of Iraq’s print heritage.
- **To develop**, through this pragmatically-sized pilot project, an approach and best practices for scanning Arabic and Middle-East language-based humanistic content, in order to facilitate access to scarcely held and disappearing materials, for a key world region.

Although today’s technology is ever-evolving, some technical practices are acknowledged as standard. The digitization of print matter as a means of creating electronic collections is evolving as an accepted methodology to preserve materials and enhance access to the digitized content.² The key element to the success of a digitizing project lies in the approach planned to manage its many components: scanning, retrieving, searching, and viewing the scanned materials. In the case of *Iraq ReCollection*, then, the project will succeed based on these critical components:

- **Digitization**: to scan a select set of Iraqi journals related to the Humanities; to share digitization tasks with and learn from the expertise of the digitization team at the Bibliotheca Alexandrina (BA), when the selected journals are published in Arabic. The BA, located in Alexandria, Egypt is the most advanced Arabic digitizing organization in the world.
- **Integration and Collaboration**: to design an electronic catalog and digital collection (available via the World Wide Web) that offers searching and display in a simple yet scalable format; to compile digitization guidelines for future scholarly digitization projects while making use of tested workflow control procedures developed at the BA.
- **Sustainability**: to configure and manage a server holding this digital collection at Yale University Library, so that the contents remain viable technically and accessible to their audience; to conduct a proof of concept of connectivity with another existing system such as OACIS; to make available the new electronic archive while designing for continuing accessibility of the archive.

In short, *Iraq ReCollection* will lay the groundwork for the creation of a digital collection of key Iraqi journals in the humanities. In turn, this effort will make possible a greater dissemination of the excellent journalistic tradition in these titles for the benefit of international research, while promoting Iraqi scholarship on a global scale. At the same time, the organized system, which will archive and make possible the display of journal content, will allow for future growth as a digital library, focusing first on Iraq but more importantly on the entire Middle East region.
Iraq ReCollection: A Proposal for Preserving Iraq’s Cultural Heritage

Project Narrative ........................................................................................................................... 37
Project Description ....................................................................................................................... 37
Project Activities .......................................................................................................................... 40
  Digitization ............................................................................................................................... 40
  Integration and Collaboration ................................................................................................. 43
  Sustainability ............................................................................................................................. 47
Methodology ................................................................................................................................. 49
  Selection criteria ...................................................................................................................... 49
  Best practices for cataloging, citations, and searching ............................................................ 49
  Hardware standardization ....................................................................................................... 50
  Software complement for digitization and Workflow control ................................................ 50
  Secure transfer of files between institutions and archive ....................................................... 52
Project Results ............................................................................................................................... 52
Project Team .................................................................................................................................. 53
Plan of Work .................................................................................................................................. 56
Budget Narrative ........................................................................................................................... 60
Appendices ........................................................................................................................................ 61
  Appendix A: Iraqi Journals covering the Humanities ............................................................. 61
  Appendix B: Curriculum vitae ................................................................................................. 61
  Appendix C: Software summary ............................................................................................... 61
  Appendix D: Hardware summary ............................................................................................. 61
  Appendix E: Summary of Guidelines developed by the BA .................................................. 62
  Appendix F: Letters of agreement to collaborate ................................................................. 64
  Appendix G: Letters of support ............................................................................................... 64
  Appendix H: OACIS Participating Institutions ...................................................................... 65
Iraq ReCollection: A Proposal for Preserving Iraq’s Cultural Heritage

Project Narrative

Project Description

Some of the most important humanistic scholarly journals titles from Iraq are held by OACIS participants. By surveying a brief selection of the Iraqi serials, one learns that these cover areas of cultural heritage such as civilization (al-Mawrid), archeology and antiquities (Sumer), history (Majallat al-mu'arrikh al-'Arab¯i), and language (Lughat al-'Arab, Majallat al-Majma‘ al-'Ilmi al-'Iraqi). It is the intention of project Iraq ReCollection to digitize significant scholarly Iraqi journals, making them readily available to researchers everywhere, in order to support the needs of the academic community within Iraq and worldwide.

Reports covering the damage to Iraq’s cultural heritage have for the most part emphasized the impact on books. For example, Shaila Dewan’s poignant account in the New York Times tells of Alia Muhammad Baker’s personal attempt to save as many books as she and her neighbors could amass before the fires took over the Central Library in Basra, Iraq. The Middle East Librarians Association Committee on Iraqi Libraries has published a report on the destruction in two Baghdad libraries: the Central library of Baghdad University, Al-Waziriyya, and the Central Awqaf library. The Middle Eastern Librarians Association (MELA) report mentions damage to 45,000 rare books and periodicals.

Monograph publications written by Iraqi authors or printed by Iraqi publishing houses are crucial to understanding the Middle East. Indeed, the Near East collection at Yale has a rich and deep assemblage of Iraqi publications in various formats. Why digitize journals then? Journals regularly offer the academic world a snapshot of discovery methods along the path traveled by
the writer-investigator. Quite often the content of a journal article covers investigations leading up to the writing of a monograph and, as such, traces the thought processes of authors or the investigations performed in order to reach conclusions published in books. Journal articles may frequently describe briefer scholarly inquiries and thinking that will never become book-length works.

The select list of journals for digitization in this project represents an important resource to academics within Iraq and here in the United States. While citation authorities such as the Web of Science do not yet include citation statistics for journals from the Middle East, academics in the U.S. have particularly highlighted the importance of journals such as *al-Mawrid Sumer*, *Lughat al-`Arab*, and *Majallat al-Majma` al-`Ilmi al-`Iraqi* when doing research. For Iraqi academics, a digital collection would help set the stage for new and necessary electronic resources. Since Iraq did not connect to the Internet till 1998, Iraqi academics have enjoyed little benefit from early electronic communication and research. However, today's efforts in the reconstruction of Iraq now include reinforcing the nation’s infrastructure to permit reliable Internet access. Therefore, our project seeks to create an electronic archive of select Iraqi journals that not only reflects the long-standing Iraqi heritage but also is available to researchers regardless of their location.

To meet the technical challenges of our project, we have envisioned collaboration based on skills, experience, and collected materials found at three institutions: Yale University Library, University of Pennsylvania Library, and the Bibliotheca Alexandrina.

Yale was among the first American colleges and universities to encourage the study of Arabic literature and Islamic culture. Edward Elbridge Salisbury was appointed professor of Arabic in 1841. Over the last 150 years, Yale has developed an extensive collection of Near
Iraq ReCollection: A Proposal for Preserving Iraq’s Cultural Heritage

Eastern library materials from three original Near East collections: the Salisbury Collection, the Landberg Collection, and the Open Arabic Collection. At present, the Near East collection is considered among the most important in this country and in the world. Yale University Library now houses more than 400,000 books relating to Near Eastern Studies and is particularly strong in classical texts, Islamic Law, History, Philosophy, and Arabic Literature. Currently the library owns more than 3,000 serials, relating to Near East studies including the major American and European scholarly journals. Over 1000 of these journals are in vernacular scripts.

In like fashion, the University of Pennsylvania has a distinguished tradition in Arabic letters. The first professorship of Arabic in the United States was established in 1782, and a professorship in Semitics was set up in 1891. By 1907 Penn offered a rich program of courses in Arabic, Aramaic, Ethiopic, Hebrew, and Syriac. Coptic and Egyptian studies were added in 1910; the Iranian languages and Turkish in 1915. The Middle East collection consists of materials in both Middle Eastern and Western languages. As of 2001, the vernacular component of the collection comprised approximately 54,500 catalogued volumes in Arabic, 9,000 volumes in Persian, 7,000 volumes in Turkish, and over 1,500 in Armenian and other area languages. The most important periodicals in Arabic, Persian and Turkish are collected.

The Bibliotheca Alexandrina (BA), the modern remaking of the great ancient library, began as a result of Egyptian governmental and UNESCO initiatives. While the BA is in its infancy in terms of collection development, its technical staff has spent the last two years developing and perfecting their digitization capabilities and is converting hundreds of books and other materials weekly for the Million Book Project and the Nasser Digital Library. Thanks to their work on projects of this technical complexity, the staff at the BA brings to our proposed
collaboration a depth of experience and expertise in Arabic text scanning that is not available currently in U.S. institutions.

Yale University Library is aware that through projects funded by the NEH and other granting agencies, the rebuilding of libraries and renovating or repairing of museums has begun. In the collaboration formed by two U.S. institutions providing the holdings and the BA affording its technical expertise, we hope to focus on reconstituting the cultural resources once held in these heritage centers by digitizing the content of selected journals and making this scholarly resource available via the Internet. Project *Iraq ReCollection* will employ accepted standards for digitizing textual materials printed in Western script and will cover new technical ground as we learn and experiment with techniques suited to the digitization of texts in Arabic. The project team in the U.S. will work in close contact with an experienced staff of digitizing experts at the BA in Egypt. In tandem, the groups will document procedures followed and lessons learned so that the project produces a scholarly resource on the Web as well as information for the benefit of other such projects to follow in the future.

**Project Activities**

**Digitization**

Because of the language challenges inherent in media printed in Arabic, it is essential that this project activity be a collaborative one. Additionally, informal scanning techniques are not appropriate for producing a quality product that has long life. We intend to divide the digitization tasks among the participants of this project according to the resources and talents available at each participating institution.
Iraq ReCollection: A Proposal for Preserving Iraq’s Cultural Heritage

First, for reasons of logistics and standardization, the actual scanning of the Iraqi journals will take place at Yale. It will be necessary to coordinate best practices between Yale and the BA so that the scanning techniques used produce consistent results. Scanning by one U.S. institution, instead of scanning at the BA, will avoid the extra costs of shipping, the coordination, and the concerns of the owning libraries. Therefore, the Yale project team will:

- Coordinate software version and functionality of scanning hardware and software used at the digitization sites for Yale and BA.⁸
- Complete the necessary training for digitization. This training will follow the guidelines drawn up by the digitizing staff at the BA whenever the materials involved are printed in Arabic. (Please see Appendix E: Summary of Guidelines developed by the BA.)
- Form a team of student workers at Yale to perform the actual scanning. Since the bulk of the serials for selection are printed in Arabic, students at Yale from the Near East Languages and Civilizations department and native speakers majoring in other disciplines will offer language skills that other labor pools do not. On other projects such as OACIS, Yale students have demonstrated themselves to be eager, reliable, and highly motivated. The OACIS student force is made up of both undergraduate and graduate students who typically spend between 4 and 6 years on campus, thus providing continued skill to the project well beyond the grant life-cycle.⁹
- Establish workflow controls, such as file naming conventions for identifying source, and set up a database for this purpose. We will also apply lessons learned from metadata implementation and best practices from the OACIS project. In the third year of the OACIS grant, we completed a digitization experiment in collaboration with the BA. In the experiment we digitized two journals in the public domain -- one at Yale, the other at
the BA – and developed a prototype digital viewer for use within the OACIS catalog.

During this experiment, we followed the BA guidelines for metadata, scanning practices, and workflow controls. ¹⁰

Next, Yale and the BA will process the scanned materials based on guidelines developed and used at the BA. This processing involves the review of all scanned images for many factors, among them: clarity, level lines of text, interference in the image caused by paper age or color. Again, by collaborating with technical experts at the participating institutions, the project team will assure that this processing phase prepares the digitized images for maximum quality results. To do this, the project team will:

• Process all scanned images, using the same software configuration at Yale and the BA for both the scanners and the processing workstations.

• Review processing team’s work for quality assurance prior to sending materials on to the Optical Character Recognition (OCR) phase. When necessary, images may be identified as requiring a new scan. Quality control procedures will be taken from a variety of sources: 1) the extensive documentation compiled by the BA, 2) standards suggested by the Research Libraries Group (RLG), and 3) pioneering work done by Cornell University. ¹¹

• Catalog the scanned pages from the journals and associated data using emerging standards for creation, description (metadata), and retrieval, such as those documented and disseminated by the Library of Congress, for example, MARC 21. ¹²
Finally, for the phase when electronic texts are converted via Optical Character Recognition (OCR) to searchable text, the participants will share technical expertise via training sessions between Yale and the BA. Conversion via OCR depends on the language of publication of the selected journal. The OCR software involved in interpreting images from a journal published in a Western script is significantly different than that used when converting one printed in Arabic into searchable text. While US libraries do have significant experience with OCR software for Western scripts, these same institutions must look to their colleagues in the Middle East for expertise when using such software to interpret Arabic text. (Please see the Methodology section below for an explanation of the OCR software.) Thus, the project team will:

- Share the processed images of the scanned pages with the digitization team at the BA, when the journal language is Arabic, to produce maximum quality processed output.
- Complete OCR interpretation for journals printed in western languages at Yale.
- Following accepted metadata standards and, using secure file transfer techniques, compile an archive from the server of all electronic files, representing the full digitization flow from scanning to OCR conversion at Yale University Library. (Please see Appendix E: Summary of Guidelines developed by the BA.)

**Integration and Collaboration**

Integration in the *Iraq ReCollection* project implies the coordination of people, expertise, and technology to produce a scholarly resource that is accessible, usable, and scalable. A technical system – in this case an electronic archive – is successful only when it addresses the needs of its target audience, drawing on the expertise and ideas of select curatorial and technical
Iraq ReCollection: A Proposal for Preserving Iraq’s Cultural Heritage

teams. The integration component stands at the center of the creative process of Project Iraq ReCollection to incorporate selection, scanning, and database design into a scalable lasting product.

During the first integration phase, the Near East curator at Yale University Library, the Near East Bibliographer at the University of Pennsylvania, and the Yale Project Manager will create a project management document from the selected titles to use as the journals proceed through the workflow. Staff from the two contributing libraries will conduct a shelf checking task that will confirm the missing issues identified in a preliminary survey, compare library catalog entries and update these based on the findings of the shelf checking. This control document will also serve in monitoring the delivery process of titles contributed to the project by the University of Pennsylvania.

Currently the document consists of 9 titles held by Yale and the University of Pennsylvania. Of these, two titles are available in both Arabic and Western languages; all others are published in Arabic. The titles are:

<table>
<thead>
<tr>
<th>Title</th>
<th>Status</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Lughat al-‘Arab</em></td>
<td>complete run</td>
</tr>
<tr>
<td><em>al-Mawrid</em></td>
<td>complete run</td>
</tr>
<tr>
<td><em>Report on the excavations in ‘Iraq</em></td>
<td>complete run</td>
</tr>
<tr>
<td><em>al-Thaqafah al-jadidah</em></td>
<td>complete run</td>
</tr>
<tr>
<td><em>Dirasat tarikhiyah</em></td>
<td>complete run</td>
</tr>
<tr>
<td><em>Sumer</em></td>
<td>1 gap issue</td>
</tr>
<tr>
<td><em>Majallat al-Majma‘ al-‘ilmi al-‘Iraqi</em></td>
<td>4 gap issues</td>
</tr>
<tr>
<td><em>Majallat al-mu‘arrikh al-‘Arabi</em></td>
<td>13 gap issues</td>
</tr>
<tr>
<td><em>al-Katib al-‘Arabi</em></td>
<td>15 gap issues</td>
</tr>
</tbody>
</table>
The table above lists the selected titles in the proposed order of digitization. The number of gap issues is noted as well. This number is based on the holdings in Sterling Memorial Library at Yale University. A survey was conducted on the selected titles. It involved a comparison of holdings shown in the OACIS union catalog, which now includes 22 member libraries. (Please see Appendix H for a listing of OACIS participating institutions.) In addition, the holdings at three non-OACIS member institutions, University of California at Berkeley, Harvard University, and University of Chicago, were reviewed. (Please see Appendix A for a detailed list of the titles showing holdings at Yale, University of Pennsylvania, other OACIS partner institutions, and these three non-OACIS members. This listing also provides details on the missing issues (including volumes and dates) and the sources in the OACIS union catalog for filling the gaps.)

It is the intention of the project team to begin with the five complete runs, which include two titles that do not present copyright problems, and *al-Mawrid*, perhaps the most important journal from Iraq for scholarly research. The journal *Sumer* has one gap issue at Yale; this issue is held by the University of Pennsylvania. The gap issues for the remaining three journals have been identified and sources for completing the runs are available within the OACIS membership. The non-OACIS libraries surveyed offer the team an extra source, if needed, to procure gap issues. Additionally, the curators will also post inquiries regarding any remaining gap issues, should this be necessary, on MELA and MELCOM listservs. In this way, we anticipate that the nine titles will be digitized in full, without missing issues.

While the digitization work proceeds on the identified complete runs, the Near East curator at Yale University Library, the Near East Bibliographer at the University of Pennsylvania, and the project’s Principal Investigator will address the copyright issues for the
remaining journals. Since the selection process has been completed as part of preparing for this grant request, the curators will be able to devote their time and effort to resolving copyright concerns.

The project team has already researched strategies regarding copyright in TITLE 17, CHAPTER 1, § 107 and 108 covering fair use and library use sections of the U.S. Copyright Act.\textsuperscript{14} We have also reviewed current Iraqi copyright law. We will exercise due diligence in contacting publishers directly and will document all contacts and attempts.\textsuperscript{15} Our intention is to deal first with academic publishers for which a contact person and address should be more easily obtained, and then with all other publishers. The project team will work with informed individuals in Iraq and the Middle East to assist in locating the extant publishers. For this, the team will seek the collaboration of USAID representatives in Iraq and the participating libraries’ book suppliers in the Middle East.\textsuperscript{16} In addition, we will send formal notice of our intentions and a full description of our project to the Iraqi embassy.

Careful attention will be paid when designing the related Web displays to note the terms and conditions of use of each digitized journal. Moreover, we will explore the use of Java applets to disable right-click options to copy or save digitized content until copyright has been cleared. Further, when adding metadata for each title, we will include in the administrative metadata all available information regarding copyright. At the same time, the project team will continue to address intellectual property concerns for the remaining journals by consulting with the Yale University counsel and a JSTOR\textsuperscript{17} copyright attorney, two vital participants in a larger Middle East digital library initiative at Yale University Library.

Next and integral to the digitization phase of the \textit{Iraq ReCollection} project is to utilize the project management document to track all phases of scanning, processing, and OCR conversion,
Iraq ReCollection: A Proposal for Preserving Iraq’s Cultural Heritage

as well as file naming and metadata implementation needed for later use in the retrieval and display of the digitized materials. In this phase, the project manager will work closely with Yale’s Near East curator and the BA digitization specialists to assure that all materials are identified and accounted for from onset to completion of the project.

Finally, along with workflow control, the team will develop a database design, implementing and refining this as digitized materials become available using the Greenstone suite of products, which has been developed based on Open Source solutions (OSS). The successful completion of this project will be marked by the ability to search, retrieve, and display archived content created from the digitization phase. The development of the archive will begin first as a prototype created early in the project lifetime. The prototype will evolve as it is modified based on design judgments resulting from periodic usability testing of the archive’s web interface.

Sustainability

The purpose of this project is to recreate in electronic form a small part of the rich Iraqi tradition seen in its serials and journals. The activities in this phase of project Iraq ReCollection will configure a server to hold the digital collection and develop and implement its connectivity to other electronic resources. We plan to preserve the content of these significant journals and provide continuing access to this content.

The server will be configured and will reside at Yale University Library. Recognizing the instability in the region, the server will remain in the U.S. The configuration of this server will follow Open Source standards. At the same time, we will install a mirror site at the BA. The maintenance of the server at Yale over the two-year period of this project will be managed
with grant funds that support current software versions, periodic updating of data, and monitoring of server activity and security.

Additionally, as the prototype system for project *Iraq ReCollection* develops, the team will focus on existing electronic resources to which *Iraq ReCollection* can be connected. The journal content emerging from our project will be integrated into Project OACIS, since the journals identified for this project are listed in OACIS. In addition, at least one of the titles held by Yale can be found in *MENAcontents*, a system of digitized Table of Contents pages developed and maintained at Universitäts-und Landesbibliothek of Sachsen-Anhalt in Halle, Germany. By connecting *Iraq ReCollection* to these two existing resources, the team will show proof of concept regarding the expansion of *Iraq ReCollection* in the future.

The project team will also focus on the technical continuity of the materials digitized, that is, in terms of the life of the storage media and the accessibility to the contents of the digital collection. This component is important so that *Iraq ReCollection* can continue to reflect the rich Iraqi writing tradition as more digitized journal content becomes available for the archive. Yale University Library is currently developing a robust digital repository to hold new materials from projects like *Iraq ReCollection* along with many other existing electronic collections. Because of this initiative designed to unify and preserve digital collections in the Yale library system, the project team for *Iraq ReCollection* will benefit from the best practices for digital sustainability as the Library’s digital repository grows. The digitized materials from *Iraq ReCollection* will have a permanent home in this emerging repository.

At the same time, the project team will contribute specialized knowledge and expertise in digitizing non-Western scripts such as Arabic. As with any granted project, it is vital to create a product that is not only scalable in design but also that can be supported and live after initial
Iraq ReCollection: A Proposal for Preserving Iraq’s Cultural Heritage

development is completed. Part of the special challenge inherent in Iraq ReCollection is the
digitization of texts printed in Arabic. Very little experience in this technical area exists today in
U.S. academic institutions. Therefore, an important sustainability task for the project team will
be to compile documentation on the successful practices used throughout the project. Most OCR
software must be instructed by the human specialist how to recognize and interpret text. It is vital
that as the OCR software is customized for languages and publishing styles, this customization is
documented and published for the use of other similar projects in the future.

Methodology

Selection criteria

The Iraqi journals held by the two contributing libraries are cataloged both in their OPAC
(On-line Public Access Catalog), as well as in OACIS. (Please see the complete list in Appendix
A.) In some cases, there is duplication of titles, that is, the same journal is held at both libraries.
This overlap will help fill in gap issues with the goal of digitizing complete runs. The titles
related to the Humanities that are shelved in Sterling Memorial Library at Yale have been
surveyed to determine an approximate page count to estimate the digitizing effort. During the
integration planning, careful notation of page counts, volume condition, and special needs will be
noted in the project management document for control purposes and, in case any specific
comment should be noted, in the metadata of the digitized journal.

Best practices for cataloging, citations, and searching

Since the selected journals will come from two American university libraries, the catalog
data in MARC 21 format, as documented by the Library of Congress, will be available for all
journals. The MARC 21 structure, Library of Congress subject headings contained in the MARC
Iraq ReCollection: A Proposal for Preserving Iraq’s Cultural Heritage

21 data, and Dublin Core standards will be used to create the database’s structure and to design the necessary search engine to gain access to the database’s digitized content. Dublin Core citations in the database records will also follow the Open Archives Initiative protocol, thus enabling OAI harvesters on the Internet to locate subjects covered in the journal content. In addition, the emerging standards from the National Information Standards Organization (NISO) will be used when creating the structure of the database and its digital objects.

Journal content will be scanned into TIFF format. TIFF is widely acknowledged as the best current format for preservation for the future. The resolution settings for the scans will depend on the OCR requirements for the language of the journal. These settings will also depend on journal-specific requirements for creating a faithful reproduction of the source material. The BA uses 300 DPI resolution for using OCR software with Arabic text, based on their standards for clarity and their requirements for storage of the digital images. We will take this into account when scanning. Once the processing and OCR interpretation is completed, PDF files will be created for easy access over the Web.

Hardware standardization

A full explanation of specific hardware components can be found in Appendix D. Both Yale and the Bibliotheca Alexandrina have experience using the Minolta PS7000 scanner. This scanner design allows for scanning of textual materials without disbinding, an important feature considering that the selected journals must remain intact for use by the patrons of their holding institutions. In addition, we will be able to use existing equipment, thus avoiding the cost of another or different scanner.

Software complement for digitization and Workflow control
Iraq ReCollection: A Proposal for Preserving Iraq’s Cultural Heritage

The most challenging element to the proposed digitization project is converting Arabic text to electronic form via Optical Character Recognition (OCR). Sakhr OCR 8, the OCR software currently used by the BA and already tested at Yale, was developed by the Sakhr Software Company in Kuwait. Currently this software is internationally recognized as the most tested and viable OCR software package for interpreting text in Arabic. This will be the primary OCR tool. In addition, the project team plans to evaluate a new product developed by NovoDynamics in Ann Arbor, Michigan. ArborScript is designed to convert paper documents or document images into computer readable form even if the original media is considered to be in a degraded form, such as smudged or low-quality copies or originals. ArborScript may make significant impact on scanning older texts that often have yellowed paper or unfamiliar fonts.

Another software complement used for this project is ScanFix from the Pegasus Imaging Corporation. ScanFix is used in the processing of images to “clean up” the image, i.e. to deskew and remove noise. In addition, ScanFix allows for the creation of small and consistent file sizes, an important factor when creating a digital collection. Descriptions of the various software packages proposed for the software complement can be found in Appendix C.

In the digitization component, the project team will follow the example of the BA team in establishing workflow control from the onset of scanning to the back-up steps for the archive being created. This workflow control is accomplished by the use of three different customized software utilities that will assure consistent file naming and metadata element assignments. In turn, from digitization and the metadata classification through to the back-up process, the project will provide excellent preservation tools for future use.

It is important also that the project team retain a strategy of flexibility as it proceeds through the technical steps of the project. For example, if CDs do not prove to be the best
method for back-up, we will readily consider other methods such as the use of hard disk storage
or portable USB drives that can be easily housed in an off-site repository as part of Yale
University Library’s long-term planning.

Secure transfer of files between institutions and archive

Collaboration implies sharing files. The challenge for the Iraq ReCollection team will be
to send files between Yale in the U.S. and the BA in Egypt both effectively and securely.
Therefore, file transfer protocols will be used to keep both the servers and their contents secure
while preserving file integrity. These transfer practices will serve as informed guidelines as
more projects involving Iraq evolve. When the quantity of data exceeds the limits of secure file
transfer protocols, we plan to use alternate methods such as the exchange of hard disks or USB
drives.

Project Results

A successful Iraq ReCollection will produce a new archive available via the Internet
based on an archive structure designed and implemented following Open Source standards and
using Open Source software for the database and the server configuration. The core list of
journals for digitization identified from the Yale University Library and the University of
Pennsylvania Library holdings will serve as the seed entry to this new archive of Iraqi
journalistic heritage.

For longevity and scalability, the best practices used in this project will be documented,
covering items such as file naming conventions, metadata entries and structure, and other
archival concerns, starting with the digitization documentation already available from the BA. In
addition, any customized OCR software drivers resulting from the digitization process will be identified and documented so that future projects can make use of these.

Dissemination of information about the new archive will occur at the outset, during and at the completion of the project. Regular announcements will appear on the home pages of the project participants’ Web sites as well as be sent to the following:

- **DigLib**: A discussion list for digital libraries researchers and librarians
- **DLF Registries**: registries of Web-searchable digital collections managed by the Digital Library Federation
- **H-Net**: Humanities & Social Sciences Online network for scholars
- **Journals including, among others**: *RLG DigiNews, Information Technology and Libraries, Journal of Arabic and Islamic Studies, Oxford Journal of Islamic Studies*
- **lis-middle-east**: email list for MELCOM-International and MELCOM-UK
- **MELANET**: A listserv devoted to Middle Eastern librarianship
- **MELANotes**: The journal of the Middle East Librarians Association
- **MESA Bulletin**: The Middle East Studies Association of North America’s journal of review
- **MESA On-Line Newsletter**: A quarterly publication on the MESA website

Participating members of OACIS will be asked to announce the new archive on their related Web sites, in particular the sites concerned with Near East studies. In addition, the connectivity between the new archive and existing electronic systems will be announced on the web sites of these systems.

**Project Team**
Iraq ReCollection: A Proposal for Preserving Iraq’s Cultural Heritage

This core team of project staff brings a wealth of experience to project Iraq ReCollection. Full CVs can be found in Appendix B.

**Principal Investigator:** Ann Okerson, Yale's Associate University Librarian with specific responsibility for Collections Development and Management, all international and area studies, and the Library’s Preservation program. She is particularly known, nationally and worldwide, for her work on the impact of electronic publishing, in the arena of electronic costs, and in copyright/licensing for the electronic environment (see <http://www.library.yale.edu/~okerson/alo.html>). In addition to the PI role, Ms. Okerson will provide support for intellectual property and economic issues.

**Project Manager and Integration Specialist:** Elizabeth A. S. Beaudin, Technical Administrator, Yale Library. Dr. Beaudin's work has included developing the software, interface, and technical communications of the recent OACIS project. Her standing as a scholar of Medieval Spanish Literature positions her ably to balance the demands of technical architecture with scholars' needs for useable content. She is also a systems architect with 20 years of information technology experience, during which she has developed both academic and administrative systems in a university environment. Dr. Beaudin will coordinate the separate components of Iraq ReCollection and will keep a day-to-day eye on progress toward meeting the specific Iraq ReCollection goals, as well as taking responsibility for accomplishing the technical achievements of the project.

**Project Outreach Director:** Simon Samoeil, Curator, Yale Library's Near East Collection and member of the Yale Council on Middle East Studies. Mr. Samoeil's specific contributions as Project Outreach Director will include evaluating the project content and establishing and shaping the partnerships needed to accomplish the work. Mr. Samoeil has a broad and deep
understanding of the field of Middle East Studies and the scholarly resources needed to support Yale's academic program. Mr. Samoeil travels regularly in the Middle East and has developed numerous contacts there over the years. Before his arrival at Yale in 1990, Mr. Samoeil worked at the libraries of Harvard and the University of Pennsylvania and at King Fahd University in Dahran, Saudi Arabia.

**Digitization Specialist**: Jennifer Weintraub, Digital Collections Specialist at Yale Library, facilitates collection development aspects of digital conversion projects at Yale and helps to coordinate all aspects (metadata, interface, preservation) of newly launched digitization projects in the Library. Ms. Weintraub led the digitization demonstration project in OACIS, and brings a solid understanding of the OCR challenges of non-Roman text to *Iraq ReCollection* from that work. Ms. Weintraub will lead Yale’s involvement in the digitization component of *Iraq ReCollection*, coordinating with the digitization partner, Bibliotheca Alexandrina.

**Information and Computer Technology Director (Middle East)**: Noha Adly, Bibliotheca Alexandrina, Technology Officer of one of the most venturesome and advanced institutions in the Middle East, where she directs many technically complex activities. Dr. Adly manages a staff of 42 specialists digitizing Arabic texts in the leading center of such expertise anywhere in the world. She will bring that expertise to the project both to support digitization projects and to provide training and consultation for the two Western institutions, seeking to develop similar competence, while carrying out the digitizing tasks. The recent addition of the BA to the Digital Library Federation membership acknowledges the technical expertise of Dr. Noha and her staff.
Iraq ReCollection: A Proposal for Preserving Iraq’s Cultural Heritage

Plan of Work

The Iraq ReCollection project will begin on January 1, 2006 and finish on December 31, 2007. During Year One of the project, a professional staff member at Yale will complete one-week digitization training already in place at the BA. At the same time, the notice seeking a group of student workers will be posted on the student employment system. Once the student team is formed, the students will be trained in digitizing methods. With the guidance of the Near East curators, the project team will prepare a project management document from the selected titles for digitizing. This document will be used to track the core data for the new archive. The digitizing team will begin scanning, processing, and OCR conversion. An integration prototype, using the Greenstone open source system, will be deployed to display materials as they are digitized. The project team will prepare press releases to announce the project and invite a review of the prototype under development by members of the academic community specializing in the Middle East.

As Year Two of the grant begins, the team will conduct a usability study of the prototype system to test the effectiveness of the search engine and the display features. Team members will incorporate modifications of the integration prototype based the results of the usability study. As the system prototype develops, integration with one other system as proof of scalability will take place. During Year Two, scanning, processing, and OCR conversion will continue and be completed. The project team will publish press releases to announce the completed project and to promote the use of the new archive. As part of the final documentation of the project, the project team will publish its findings on the creation of a digital collection and on long term funding.
Iraq ReCollection: A Proposal for Preserving Iraq’s Cultural Heritage

Yale University Library is requesting funding for 1.5 FTE equivalents of student workers whose total effort will be devoted to the proposed project. These students will be trained in house by professional staff at Yale. The team will perform all tasks related to the workflow: scanning, processing, and OCR. Their work will be supervised and reviewed on a regular basis by the project manager, the digitization specialist, and the BA technical consultants.
Both Yale and Penn libraries will contribute varying amounts of effort of the individuals mentioned as follows.

Ann Okerson, AUL for Collection Development at Yale University Library, will serve PI for this project and oversee intellectual property and economic issues.

Elizabeth Beaudin, Technical Administrator for the OACIS project at Yale University Library, will serve as project manager and will coordinate and guide the activities in all components of the project.

Simon Samoeil, Near East Curator at Yale University Library, will evaluate the project content and guide the partnerships needed to accomplish the work.

Jennifer Weintraub, Digital Collections Specialist at Yale University Library, will oversee the digitization component of the project.

Roberta Pilette, Head of the Preservation Department, will assist in consultation and support of the digitization and metadata components.

William J. Kopycki, Middle East Bibliographer at the Van Pelt-Dietrich Library, University of Pennsylvania, will work in close cooperation with Mr. Samoeil to evaluate the digitized content.
Iraq ReCollection: A Proposal for Preserving Iraq’s Cultural Heritage

Plan of Work timeline

<table>
<thead>
<tr>
<th>NEH Year</th>
<th>1</th>
<th>2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Calendar Year</td>
<td>'06</td>
<td>'07</td>
</tr>
<tr>
<td>Calendar Quarter</td>
<td>1</td>
<td>2</td>
</tr>
</tbody>
</table>

**Digitization**
- System configuration
- Student team – job notice
- Staff training at BA
- Student team - training
- Metadata planning
- Scanning
- OCR processing
- Metadata implementation

**Integration**
- Selection & Analysis
- Copyright documentation
- Prototype interface
- User interface
- Workflow procedures
- Database implementation
- Telecommunications
- Testing, Documentation

**Longevity**
- Server configuration
- Connect to OACIS
- Connect to MENALib
- Sustainability planning
- Publication of practices
- Press releases

**Evaluations**
- Prototype interface
- User interface
- Content retrieval
- Digitization usability
- Usage statistics
Budget Narrative

1. Digitization specialist: a contractual consulting fee is added to the budget to cover the exchange of expertise between the Bibliotheca Alexandrina and Yale. As explained in the work plan, all scanning will take place at Yale; since the BA has the expertise in OCR processing, scanned images will be sent to the BA for this step.

2. Travel:
   a. Middle East Digitization training: We plan to send 1 individual on staff at Yale for intensive training in scanning and workflow control to the BA. This item covers travel to BA-Alexandria, 2 week stay, Round Trip air fare, and expenses.
   b. Middle East Technical site visit: We plan to send 1 individual on staff at Yale to the BA at the end of Year One to conduct an assessment of the workflow controls and the OCR process. These funds cover travel to BA-Alexandria, 1 week stay, Round Trip air fare, and expenses.

3. Student workers at Yale currently earn $12/hr. Therefore, for the first week we estimate 40 hrs/wk x 48 wks or 1 FTE equivalent for the 1st yr only and plan to increase this to 1.5 FTE for the 2nd year.

4. Technical training: These funds are projected for completing a course in XML (Extensible Markup Language) which is the mark-up language to create documents that are self-describing. This mark-up language is used when dealing with databases holding digital content and assists in generating display pages from the metadata descriptions of digital content.
Appendices

Appendix A: Iraqi Journals covering the Humanities

Appendix B: Curriculum vitae

Ann Okerson
Simon Samoeil
Elizabeth A. S. Beaudin
Jennifer Weintraub

Appendix C: Software summary

<table>
<thead>
<tr>
<th>Name</th>
<th>Purpose</th>
<th>Version</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sakhr Software</td>
<td>OCR of Arabic text</td>
<td>OCR 8 Gold</td>
</tr>
<tr>
<td>Adobe Acrobat</td>
<td>Creation of PDF files</td>
<td></td>
</tr>
<tr>
<td>AcdSee</td>
<td>Processing of scanned images</td>
<td>8</td>
</tr>
<tr>
<td>Linux Enterprise</td>
<td>Server software</td>
<td>AS 4</td>
</tr>
<tr>
<td>Greenstone</td>
<td>OSS suite of utilities including basic database structure</td>
<td></td>
</tr>
<tr>
<td>ABBYY FineReader</td>
<td>Processing of scanned images</td>
<td>OCR 7.0</td>
</tr>
<tr>
<td>SRZ ProScan Book</td>
<td>Scanning software installed for Minolta 7500 scanner</td>
<td>V2.1</td>
</tr>
<tr>
<td>ScanFix</td>
<td>Processing of TIFF images</td>
<td></td>
</tr>
</tbody>
</table>

Appendix D: Hardware summary

<table>
<thead>
<tr>
<th>Name</th>
<th>Purpose</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Minolta 7500 scanner</td>
<td>Scan and create images of text</td>
<td>Already owned by Yale</td>
</tr>
<tr>
<td>DELL Optiplex workstation</td>
<td>Processing and OCR of images</td>
<td>To be purchased with grant funds</td>
</tr>
</tbody>
</table>
The BA Digital Lab has compiled a 71-page manual on scanning and processing in which quality control standards for each digitization step have been established based on their experience with the Million Book project. This will be the initial working document for the proposed project. At the BA, the workflow and best practices covered in this guide fall into three categories: scanning, processing, and Optical Character Recognition (OCR).

- **Scanning** – the physical image of the scanned page is created by using a flat bed scanner, currently the Minolta 7500. The image is saved in TIFF format in a directory on a connected PC workstation using naming conventions that identify the title, dates and place of publication, and language of content. Scanning is done at 300 DPI (dots per inch, also referred to as PPI, pixels per inch) based on the experience of the BA.

- **Processing** – during this phase, the TIFF image is reviewed for clarity, level lines of text, and interference in the image caused by paper age or color. The digitizing specialist sets parameters to “clean” the image, e.g. to de-speckle, that is to remove stray marks in the image that would interfere with the OCR step. The TIFF image is saved after processing into PDF format when the Arabic text is considered searchable. Depending on the age of the journal and the font used, a text in Arabic may be deemed “unsearchable” or not worthy of the OCR step because the OCR quality of certain fonts falls well below the expected quality percentiles in the BA guidelines. When this is the case, the image will be saved in JPEG 2000 format for display purposes only.

- **OCR** – special software is used to interpret the PDF file so that the black and white image of text is turned into its machine-readable equivalent, i.e. into a numerical encoding.
scheme that permits the text and the words in the text to be searched. At this stage a TXT file is created containing the searchable text. This TXT file is stored with its corresponding PDF file in what is referred to as a “text behind image” approach. In other words, as the patron views the PDF file, searches performed by the patron are processed against the contents of the TXT file which remains hidden to the patron.

To increase quality, specialists are trained in all phases of the digitizing workflow. Specialists shift from one task to another, for example processing the TIFF images created by another specialist, so that all collaborate in the finished product.
Appendix F: Letters of agreement to collaborate

Noha Adly, ICT/ISIS Director, Bibliotheca Alexandrina, Alexandria, Egypt

H. Carton Rogers, Vice Provost and Director of Libraries, University of Pennsylvania

Heiner Schnelling, Director, Universitäts-und Landesbibliothek of Sachsen-Anhalt in Halle, Germany

Appendix G: Letters of support

Mohammed Alwan, Lecturer in Arabic, Department of German, Russian and Asian Languages & Literatures, Tufts University

Sinan Antoon, Asst. Professor, Asian and Middle Eastern Languages and Cultures, Dartmouth College

Paul Auchterlonie, Librarian for Middle East Studies & Chair MELCOM (UK), University of Exeter, UK

Paul Conway, Director, Information Technology Services, Duke University Libraries

Benjamin R. Foster, William M. Laffan Professor of Assyriology and Babylonian Literature, Department of Near Eastern Languages and Cultures, Yale University

Beatrice Gruendler, Professor and Chair, Department of Near Eastern Languages and Cultures, Yale University

Jonathan Rodgers, Head, Near East Division, University of Michigan Library

Mary St. Germain, Head, Near East Section, University of Washington Libraries

Arnoud J. M. Vrolijk, Chair, MELCOM International and Curator of Middle East Collections, Leiden University Library, The Netherlands
## Appendix H: OACIS Participating Institutions

<table>
<thead>
<tr>
<th>#</th>
<th>Institution</th>
<th>Country</th>
<th>Remarks</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>American University of Beirut</td>
<td>Lebanon</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Bibliotheca Alexandrina</td>
<td>Egypt</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>Cornell University</td>
<td>USA</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>University of Michigan</td>
<td>USA</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>New York Public Library</td>
<td>USA</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>New York University</td>
<td>USA</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>Ohio State University</td>
<td>USA</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>University of Pennsylvania</td>
<td>USA</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>Princeton University</td>
<td>USA</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>School of Oriental and African Studies (SOAS)</td>
<td>United Kingdom</td>
<td>Test dataset under review</td>
</tr>
<tr>
<td>11</td>
<td>Stanford University</td>
<td>USA</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>University of Texas at Austin</td>
<td>USA</td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>ULB Sachsen-Anhalt, Halle</td>
<td>Germany</td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>Tishreen University</td>
<td>Syria</td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>University of Arizona</td>
<td>USA</td>
<td>Method of extract under discussion</td>
</tr>
<tr>
<td>16</td>
<td>University of Balamand</td>
<td>USA</td>
<td>Methods of delivery under discussion</td>
</tr>
<tr>
<td>17</td>
<td>University of California at Los Angeles</td>
<td>USA</td>
<td>Method of extract under development</td>
</tr>
<tr>
<td>18</td>
<td>University of Illinois at Urbana-Champaign</td>
<td>USA</td>
<td></td>
</tr>
<tr>
<td>19</td>
<td>University of Jordan</td>
<td>Jordan</td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>University of Utah</td>
<td>USA</td>
<td>Test dataset under review</td>
</tr>
<tr>
<td>21</td>
<td>University of Washington</td>
<td>USA</td>
<td></td>
</tr>
<tr>
<td>22</td>
<td>Yale University</td>
<td>USA</td>
<td></td>
</tr>
<tr>
<td>23</td>
<td>Yale Law School</td>
<td>USA</td>
<td></td>
</tr>
</tbody>
</table>

*As of August 1, 2005*
Evidence of the growing interest in creating digital collections is seen in the increasing number of complex and successful ventures. From this work, a rich corpus of documentation has emerged for the benefit of libraries, museums, and other cultural heritage institutions that are considering digitization projects. See, for example, the publication from the New England Document Conservation Center (NEDCC) entitled “Handbook for Digital Projects: A Management Tool for Preservation and Access” at: <http://www.loc.gov/marc/marc.html>. The Library of Congress also documents standards on metadata at: <http://memory.loc.gov/ammem/collections/habs_haer/hhdigit.html>. Also, please see: <http://www.loc.gov/standards/metadata.html> for basic metadata standards which will be used in conjunction with the best practices used by the BA to develop an evolving set of documentation detailing the digital collection creation. 

Our initial set of standards will combine the practices documented at the BA (Please see the Appendix for a summary) along with those used by various institutions including the Library of Congress, the Digital Library Federation (DLF), Cornell University, Harvard University, and the New England Document Conversion Center (NEDCC). These and other institutions and projects are mentioned throughout the proposal in specific context. An example of coordinating the configurations at the two sites might include verifying that all scanners to be used include a gray-scale chip, i.e. that all scanners are functioning with the same version of software and the same hardware capabilities.

Although a student team may not appear to provide the continuity needed, we have found from other projects that the students welcome challenging work and stay longer in such positions. In a different project, the Colorado Digitization Project recruited volunteers from the community to form a work force. Continuity was a key factor and was managed by a training program that all workers completed prior to joining the team. In addition, workers had documentation at hand for easy reference. We find this practice to suit our needs and therefore we will incorporate a training program and standard documentation for all students hired.

The BA Digital Lab has compiled a 71-page guide on scanning and processing in which quality control standards for each digitization step are established. While this will be the primary document for the project, our team will also refer to accepted practices noted by U.S. libraries and research institutions with significant preservation experience. For the RLG site, please see: <http://www.rlg.org/index.php>. Cornell provides extensive documentation on quality control at: <http://www.library.cornell.edu/iris/dpo/prespubs.html>. According to the definition provided by the Library of Congress, “The MARC formats are standards for the representation and communication of bibliographic and related information in machine-readable form.” Please see <http://www.loc.gov/marc/marc.html>. The Library of Congress also documents standards on metadata at: <http://www.loc.gov/standards/mods/>.

MELA is the Middle Eastern Librarian Association and MELCOM is its European counterpart. Both professional associations maintain very active communication via listservs. For MELA, see: <http://www.mela.us/>; for MELCOM, please see: <http://www.uni-bamberg.de/unibib/melcom/home.html>.

We have also consulted other digital projects for information regarding copyright such as the Colorado Digitization Project. Please see their organized resources at: <http://www.cdpheritage.org/resource/legal/resource.cfm>.

We also plan to make use of the Indiana University copyright checklist as part of our documentation of all copyright inquiries. Please see: <http://www.copyright.iupui.edu/checklist.htm>.

USAID has made significant positive impact in many Iraqi sectors. We plan to contact those USAID representatives working on educational projects such as the installation of a computer lab at the University of Mosul.
Iraq ReCollection: A Proposal for Preserving Iraq’s Cultural Heritage

Yale University Library has benefited from an excellent and long-term collaboration with Leila Books in Cairo, Egypt. Mr. George Fawzi, the owner and CEO, can be called on to contact publishers in the area.

17 JSTOR (http://www.jstor.org/) is a scholarly journal archive offering scanned images of original scholarly publications.

15 The Open Source movement has evolved over the last twenty years first among individual software developers who shared program code and design ideas. It is now organized by an international initiative. See <http://www.opensource.org/docs/definition.php> for more on accepted Open Source standards. And in particular, for a description of the software and utilities, please see: <http://www.greenstone.org>. Based on our experience with the OACIS project, we believe the Open Source approach will provide a scalable and sustainable design.

For example, instead of using licensed server software from Microsoft or Sun Systems, an Open Source server approach would use the Linux operating system software for the configuration of the server.

19 MENAContents developed by the Universitäts-und Landesbibliothek of Sachsen-Anhalt in Halle, Germany contains digitized tables of contents for approximately 200 journal titles related to the Middle East. Please see: <http://www.bibliothek.uni-halle.de/ssg/inhalt.htm>.

21 The Open Archives Initiative (OAI) seeks to facilitate interoperability in the dissemination of content over the Internet. Harvesters are software programs that search the Internet for metadata conforming to published OAI standards. For more information, see: <http://www.openarchives.org/documents/index.html>.


23 To determine the level of scanning that will provide a faithful reproduction, we will use such established guidelines as the Digital Library Federation's Benchmark for Faithful Reproduction of Monographs and Serials (http://www.diglib.org/standards/bmarkfin.htm) and NARA Technical Guidelines for Digitizing Archival Material for Electronic Access (http://www.archives.gov/research_room/arc/arc_info/guidelines_for_digitizing_archival_materials.html). Other resources providing useful guidance on scanning practices include the Library of Congress and the Colorado Digitization Project.


25 Please see: http://www.pegasusimaging.com/scanfix.htm